The ISME Journal (2009) 3, 31–46
© 2009 International Society for Microbial Ecology All rights reserved 1751-7362/09 $32.00
www.nature.com/ismej

ORIGINAL ARTICLE

Phenotypic and genetic diversification of *Pseudanabaena* spp. (cyanobacteria)

Silvia G Acinas[1,3,4], Thomas HA Haverkamp[1,4], Jef Huisman[2] and Lucas J Stal[1,2]
[1]*Department of Marine Microbiology, Netherlands Institute of Ecology, NIOO-KNAW, AC Yerseke, The Netherlands and* [2]*Department of Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*

***Pseudanabaena* species are poorly known filamentous bloom-forming cyanobacteria closely related to *Limnothrix*. We isolated 28 *Pseudanabaena* strains from the Baltic Sea (BS) and the Albufera de Valencia (AV; Spain). By combining phenotypic and genotypic approaches, the phylogeny, diversity and evolutionary diversification of these isolates were explored. Analysis of the *in vivo* absorption spectra of the *Pseudanabaena* strains revealed two coexisting pigmentation phenotypes: (i) phycocyanin-rich (PC-rich) strains and (ii) strains containing both PC and phycoerythrin (PE). Strains of the latter phenotype were all capable of complementary chromatic adaptation (CCA). About 65 kb of the *Pseudanabaena* genomes were sequenced through a multilocus sequencing approach including the sequencing of the16 and 23S rRNA genes, the ribosomal intergenic spacer (IGS), internal transcribed spacer 1 (ITS-1), the *cpcBA* operon encoding PC and the IGS between *cpcA* and *cpcB*. In addition, the presence of *nifH*, one of the structural genes of nitrogenase, was investigated. Sequence analysis of ITS and *cpcBA*-IGS allowed the differentiation between *Pseudanabaena* isolates exhibiting high levels of microdiversity. This multilocus sequencing approach revealed specific clusters for the BS, the AV and a mixed cluster with strains from both ecosystems. The latter comprised exclusively CCA phenotypes. The phylogenies of the 16 and 23S rRNA genes are consistent, but analysis of other loci indicated the loss of substructure, suggesting that the recombination between these loci has occurred. Our preliminary results on population genetic analyses of the PC genes suggest an evolutionary diversification of *Pseudanabaena* through purifying selection.**
The ISME Journal (2009) **3,** 31–46; doi:10.1038/ismej.2008.78; published online 4 September 2008
Subject Category: microbial population and community ecology
Keywords: *Pseudanabaena*; cyanobacteria; complementary chromatic adaptation; microbial diversity; phylogeny; evolutionary diversification

Introduction

Cyanobacteria are the dominant component of the phytoplankton in many freshwater and marine environments where they may form nuisance blooms (Chorus and Bartram, 1999; Huisman *et al.*, 2005; Granéli and Turner, 2006). The attention is usually toward the larger species that form aggregates and possess gas vesicles that make them buoyant and therefore accumulate at the surface. These species fix N_2 and may be toxic and may cause serious environmental and socioeconomical problems. However, often it is not recognized that smaller cyanobacteria exceed the larger species in terms of biomass and activity and therefore play a major role in the ecosystem dynamics. Several reports mentioned the occurrence of the tiny filamentous *Pseudanabaena* in cyanobacterial blooms in brackish and freshwater ecosystems (Vasconcelos and Pereira, 2001; Stal *et al.*, 2003; Gkelis *et al.*, 2005; Zwart *et al.*, 2005; Kim *et al.*, 2006; Willame *et al.*, 2006). Nevertheless, *Pseudanabaena* spp. has only had little attention in the scientific community where limited data exist about their phenotypic and genetic characteristics.

Pseudanabaena species are nonheterocystous cyanobacteria belonging to the order of Oscillatoriales. The family of *Pseudanabaenaceae* is characterized by simple trichomes with a width less than 4 μm. The cells are longer than wide, possess parietal thylakoids, contain polar gas vesicles and the cross walls are conspicuously constricted (Castenholz *et al.*, 2001; Komárek, 2003). Some strains display complementary chromatic adaptation (CCA). This process allows these organisms to regulate the ratio

Correspondence: LJ Stal, Department of Marine Microbiology, Netherlands Institute of Ecology, NIOO-KNAW, PO Box 140, Yerseke, 4400 AC, The Netherlands.
E-mail: l.Stal@nioo.knaw.nl
[3]Current address: Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar, CSIC, 08003, Barcelona, Spain.
[4]These authors contributed equally to this work.
Received 21 May 2008; revised 10 July 2008; accepted 17 July 2008; published online 4 September 2008

of the accessory photosynthetic pigments phycocyanin (PC) and phycoerythrin (PE), which helps them to adapt to the prevailing light spectrum (reviewed by Kehoe and Gutu (2006)) thereby favoring their persistence in competition against other species (Stomp *et al*., 2004, 2008). Most cultured strains reveal gliding motility and some are capable of anaerobic N_2 fixation (Rippka and Herdman, 1992).

Morphologically, *Pseudanabaena* resembles *Limnothrix* making their identification difficult. The main differences are the somewhat wider cells (1–6 μm) and the less distinct constriction of the cross walls in *Limnothrix* (Castenholz *et al*., 2001). Although they are rarely recognized as dominant organisms, *Pseudanabaena* as well as *Limnothrix* species occur and form blooms in eutrophic water bodies and occasionally dominate the phytoplankton (Mayer *et al*., 1997; Rücker *et al*., 1997; Zwart *et al*., 2005). *Limnothrix* is typically found in meso- to eutrophic freshwater ecosystems, whereas *Pseudanabaena* is more widely distributed and occurs in diverse aquatic as well as in benthic environments (Castenholz *et al*., 2001; Zwart *et al*., 2005; Diez *et al*., 2007). On the basis of the 16S rRNA gene, *Pseudanabaena* and *Limnothrix* form a monophyletic cluster within the cyanobacteria (Zwart *et al*., 2005; Willame *et al*., 2006). However, the available molecular data of the *Pseudanabaena*/*Limnothrix* group is scarce and consists of a limited number of environmental sequences and a few isolates. This limited data does not resolve the phylogeny of the *Pseudanabaena*/*Limnothrix* group.

Here, we present a multiphasic phenotypic and genotypic approach to explore the diversity, phylogeny and evolutionary diversification of *Pseudanabaena* strains isolated from two distant geographical regions in Europe, the BS, a large brackish basin in the North and Albufera de Valencia (AV), a coastal lagoon in the South. The BS is one of the largest bodies of brackish water in the world. It is a eutrophic system that develops blooms of the conspicuous N_2-fixing heterocystous cyanobacteria *Aphanizomenon* and *Nodularia*, of which the latter is toxic (Stal *et al*., 2003). However, the dominant component of the cyanobacterial community in the BS consists of a colorful mixture of unicellular picocyanobacteria of the *Synechococcus* group and the tiny filamentous *Pseudanabaena* (Stal *et al*., 2003; Stomp *et al*., 2007; Haverkamp *et al*., 2008). The AV is a highly eutrophic coastal freshwater lagoon in Spain that is fed by streams, rivers and irrigation channels carrying fertilizer from the surrounding rice fields. The Albufera is characterized by dense water blooms of cyanobacteria among which *Pseudanabaena* spp. is a dominant group of organisms (Romo and Miracle, 1994; Villena and Romo, 2003). This report describes the phenotypic and genetic diversification found within *Pseudanabaena* from two geographical locations and it is thereby a first step toward an understanding of the ecology of this interesting but poorly known filamentous bloom-forming cyanobacteria.

Materials and methods

Isolation, cultivation and strain collection

The *Pseudanabaena* strains used in this study were isolated between 1995 and 2004 (Table 1). Strain CCY9508 was isolated in 1995 from the BS (Bornholm Sea). The Spanish strains were isolated in 1997 from AV. Although we used the same basic medium for the isolation of the cyanobacteria, we varied the source of nitrogen and applied two different approaches: (i) size fractionation filtrations and (ii) dilution to extinction. For the isolation of these strains, water was prefiltered through 2 μm mesh plankton net. Subsequently, the filtrate was spread onto 0.7% agarose medium in Petri dishes. Strain CCY9508 was isolated on a mixture consisting of 1/3 volume ASNIII + 2/3 volume BG11 medium with a salinity of 12.2‰ (Rippka *et al*., 1979). The Spanish isolates were isolated on the freshwater BG11 medium using nitrate as nitrogen source and incubation at 14 °C. Single colonies were picked from the agarose plates and repeatedly transferred until axenic monoclonal strains were obtained. The other BS strains were isolated from samples collected at various stations in the Gulf of Finland (from 59.1 °N 22.2 °E to 60.0 °N 26.2 °E) during a research cruise in July 2004. Water samples were collected from defined depths using a rosette sampler. *Pseudanabaena* strains were isolated using two different approaches. In one approach, water samples were fractionated, first using one layer and subsequently two layers of plankton net (20 μm mesh) under gentle vacuum. This filtrate was successively filtered through 5 μm, 1 μm and 0.45 μm membrane filters. Finally, the 0.45 μm filter was transferred to a sterile 10 cm Petri dish filled with a mixture of 4/5 parts BG11 and 1/5 parts ASNIII medium, containing NH_4Cl ($0.05\ g\,l^{-1}$) as the nitrogen source. In the other approach, water was filtered through 20 μm plankton net and subsequently diluted to extinction in 96 deep-well microtiterplates (Nunc Inc., Thermo Fischer Scientific, Langenselbold, Germany) containing a mixture of 4/5 parts BG11 and 1/5 parts ASNIII medium containing NH_4Cl ($0.05\ g\,l^{-1}$). In both approaches, the cells were first grown under a light regime of 10 μmol photons $m^{-2}\,s^{-1}$ for 2 weeks at 20 °C. Subsequently, the light intensity was increased to 20 μmol photons per m^2 per second and the cultures were incubated for another 6 weeks at the same temperature. Trichomes growing on the filters or at the surface of the wells were picked and transferred to solid media. To obtain monoclonal axenic strains, trichomes were repeatedly transferred. Once obtained, monoclonal axenic strains were maintained in their specific growth medium in the Culture Collection Yerseke (CCY; Table 1).

Table 1 Characteristics of *Pseudanabaena* strains used in this study, year and location of isolation and PCR amplification for *nifH* genes

Strain	*Isolation details*		*Growth*	*Morphological characteristics*		*Motility*[c]	*CCA*	*Absorption ratio*			*nifH PCR Amplification*
	Origin	*Year*	*Medium*[a]	*Cell size (μm)*[b]	*PGV*			*570/625*[d]	*680/570*[e]	*680/625*[f]	
CCY9701	Albufera de Valencia	1997	BG 11	6.1 × 1.5	+	–	–	0.75	1.22	0.91	+
CCY9702	Albufera de Valencia	1997	BG 11	6.7 × 1.4	+	–	ND	ND	ND	ND	+
CCY9703	Albufera de Valencia	1997	BG 11	5.5 × 1.5	+	+	+	1.07	0.9	0.96	ND
CCY9704	Albufera de Valencia	1997	BG 11	6.0 × 1.6	+	–	+	1.02	1.07	1.1	ND
CCY9705	Albufera de Valencia	1997	BG 11	4.3 × 1.6	+	+	+	1.18	0.96	1.14	ND
CCY9709	Albufera de Valencia	1997	BG 11	4.7 × 1.5	+	+	+	1.06	1.01	1.08	ND
CCY9710	Albufera de Valencia	1997	BG 11	3.5 × 1.4	–	+	–	0.82	1.21	0.99	ND
CCY9712	Albufera de Valencia	1997	BG 11	4.9 × 1.5	+	+	–	0.67	1.22	0.81	ND
CCY9714	Albufera de Valencia	1997	BG 11	4.9 × 2.0	+	+	–	0.69	1.41	0.97	ND
CCY9715	Albufera de Valencia	1997	BG 11	4.9 × 1.4	+	–	+	1.18	1.08	1.27	ND
CCY9508	Baltic Sea	1995	1/3A+2/3B	6.1 × 1.5	+	+	+	0.67	1.51	1.01	ND
CCY0471	Baltic Sea	2004	1/5A+4/5B	3.2 × 1.4	+	+	+	1.04	1.48	1.53	+
CCY0472	Baltic Sea	2004	1/5A+4/5B	5.2 × 1.2	+	+	+	1.17	1.52	1.78	+
CCY0473	Baltic Sea	2004	1/5A+4/5B	2.5 × 1.5	+	+	+	1.28	1.12	1.43	+
CCY0474	Baltic Sea	2004	1/5A+4/5B	2.8 × 1.5	+	+	+	1.49	1.18	1.75	–
CCY0475	Baltic Sea	2004	1/5A+4/5B	2.9 × 2.0	+	+	+	1.52	0.88	1.34	–
CCY0476	Baltic Sea	2004	1/5A+4/5B	2.2 × 1.6	+	+	–	0.62	1.45	0.89	–
CCY0477	Baltic Sea	2004	1/5A+4/5B	2.6 × 2.0	+	+	+	1.29	0.95	1.23	+
CCY0478	Baltic Sea	2004	1/5A+4/5B	2.7 × 1.6	+	+	+	1.21	1.08	1.31	–
CCY0479	Baltic Sea	2004	1/5A+4/5B	2.6 × 1.9	+	+	+	1.4	0.92	1.28	–
CCY0480	Baltic Sea	2004	1/5A+4/5B	2.4 × 1.3	+	+	–	0.74	1.42	1.05	–
CCY0481	Baltic Sea	2004	1/5A+4/5B	2.1 × 1.9	+	+	+	1.52	0.96	1.47	–
CCY0482	Baltic Sea	2004	1/5A+4/5B	2.3 × 1.2	+	+	–	0.55	1.92	1.06	–
CCY0483	Baltic Sea	2004	1/5A+4/5B	2.4 × 1.9	+	+	+	1.46	0.89	1.31	–
CCY0484	Baltic Sea	2004	1/5A+4/5B	2.4 × 1.9	+	+	–	0.6	1.91	1.14	–
CCY0485	Baltic Sea	2004	1/5A+4/5B	2.3 × 1.3	+	+	–	0.59	2.51	1.5	–
CCY0486	Baltic Sea	2004	1/5A+4/5B	2.6 × 1.5	+	+	+	1.17	1.26	1.47	–
CCY0488	Baltic Sea	2004	1/5A+4/5B	3.0 × 1.3	+	+	–	0.69	1.62	1.11	–

Abbreviations: CCA, complementary chromatic adaptation; Chl, chlorophyll; ND, not determined; PGV, polar gas vesicles.
[a]Growth media description can be found in Rippka *et al*., 1979 (1/3A+2/3B: 1 part ASNIII medium+2 parts BG11 medium).
[b]Cell size based on the average length × width of at least 40 cells.
[c]Motility on agarose plates.
[d]Ratio between absorption at 570 nm (phycoerythrin) and 625 nm (phycocyanin).
[e]Ratio between absorption at 680 nm (Chl*a*) and 570 nm (phycoerythrin).
[f]Ratio between absorption at 680 nm (Chl*a*) and 625 nm (phycocyanin).

Morphology and microscopy
Cells were collected from exponential or stationary liquid cultures and were fixed in a mixture of 1% (w/v) formaldehyde and 0.05% (w/v) glutaraldehyde and subsequently stored at −80 °C until they were analyzed (Biegala *et al.*, 2003). Microscope slides were prepared by covering them with a thin layer of 1% (w/v) molten agarose (50 °C; Sigma-Aldrich, Zwijndrecht, The Netherlands) that was allowed to solidify shortly before cells were applied. The slides were examined using a Zeiss Axiophot microscope equipped with a ProgRes C10 plus digital imaging system (JENOPTIK Laser, Optik, Systeme GmbH). The images were subsequently processed using ProgRes CapturePro2.0 software (JENOPTIK Laser, Optik, Systeme GmbH). From at least 30 cells in each culture, the width and length were measured.

In vivo *absorption spectra*
Exponential or stationary liquid cultures grown under white light (20 μmol photons per m^2 per second) were used for the determination of the *in vivo* absorption spectra. Spectra were measured from 400 to 750 nm using a Varian Cary 100 Bio equipped with an integrating sphere DRA-CA-3300. Distilled water was used as reference.

Determination of CCA
To test for the capacity of CCA, the strains were cultured on solid (agarose) media. Each strain was inoculated in two Petri dishes (Greiner Bio-One) that were incubated under a different color of light. Green light was obtained using Lee filter no. 124 (dark green) and red light through Lee filter no. 26 (red). The incident white light intensity was 100 μmol photons per m^2 per second. To document CCA, the cultures were photographed after 2 weeks of growth under monochromatic light. Subsequently, the cultures were changed to the other color of monochromatic light and incubated another 2 weeks after which they were documented again. A change from green to reddish/black phenotype and *vice versa*, was taken as evidence for CCA.

The strains CCY9703 and CCY9710 were also cultured in liquid medium. Three liquid cultures of each strain were grown under white light conditions (14 μmol photons per m^2 per second) until stationary phase was reached. Subsequently, the three cultures were grown under white, green or red light. After 2 weeks, the cultures were sampled and the *in vivo* absorption spectra were measured as described above.

DNA isolation
For DNA extraction, cells were collected from exponentially growing or stationary phase cultures. Briefly, two milliliters of culture was centrifuged in a tabletop centrifuge (Eppendorf type 5424) at 10000 r.p.m. for 1 min at room temperature. The supernatant was removed and the cell pellet was resuspended in lysis buffer provided by the Powersoil DNA extraction kit (MoBio Laboratories, Inc., Carlsbad, CA, USA). DNA extraction was performed following the instructions of the manufacturer. DNA quantity and quality were checked by running agarose gels as well as spectrophotometrically (Nanodrop ND1000).

PCR and sequencing
PCR reactions were performed using the PCR primers listed in Table 2. The B1055F and PitsE-cyanR were used to amplify the end of the 16S rRNA gene plus the internal transcribed spacer 1 (ITS-1). The 23S rRNA gene was amplified with 129F and 2241R primers and the PC operon (*cpcBA*) with *cpcA*R and *cpcB*F primers. Finally, *nifH* PCR amplification were performed by the set of primers

Table 2 List of the primers used for PCR and sequencing in this study

Primer name	*Target*	*Sequence (5′-3′)*	*Reference*
Bact1055F[a]	16S rRNA gene	AATGGCTGTCGTCAGCTCGT	Garcia-Martinez *et al.* (1999)
PitsE-cyanR[a]	23S rRNA gene	CTCTGTGTGCCAAGGTATC	Ernst *et al.* (2003)
129F[b]	23S rRNA gene	CYGAATGGGRVAACC	Hunt *et al.* (2006)
2241R[b]	23S rRNA gene	ACCGCCCCAGTHAAACT	Lane (1991)[c]
CpcAR[d]	Phycocyanic operon	TAGTGTAAAACGACGGCCAGT TGYYTKCGCGACATGGA	Robertson *et al.* (2001)
CpcBF[d]	Phycocyanic operon	TAGCAGGAAACAGCTATGACG TGGTGTARGGGAAYTT	Robertson *et al.* (2001)
nifH1	nifH	TGYGAYCCNAARGCNGA	Zani *et al.* (2000)
nifH2	nifH	ADNGCCATCATYTCNCC	Zani *et al.* (2000)
nifH3	NifH (internal primer)	ATRTTRTTNGCNGCRTA	Zani *et al.* (2000)
nifH4	NifH (internal primer)	TTYTAYGGNAARGGNGG	Zani *et al.* (2000)

[a]These primers were also used to obtain the complete ITS-1 sequence.
[b]These primers were used to amplify the complete 23S rRNA gene from the *Pseudanabaena* strains.
[c]This primer was revised recently by Hunt *et al.*, 2006.[50]
[d]These primers were used to amplify the subunits β and α of the phycocyanin operon plus the intergenic spacer (IGS) between both genes.

described in Table 2. Each reaction contained 0.2 mM of each dNTP, 2 mM $MgCl_2$, 5 or 10 pmol of each primer, 1 µl template DNA (5–10 ng^{-1}), 1 × PCR buffer and 1 U HotStarTaq (Qiagen GmbH, Hilden, Germany). MQ-grade H_2O was added to a final volume of 30 µl. The PCR reactions were carried out in a GeneAmp System 2700 thermocycler (Applied Biosystems, Foster City, CA, USA). The PCR program was used to amplify the 16S rRNA-ITS-1 region, the 23S rRNA gene and the *cpcBA* operon were as follows: a hot start at 94 °C of 15 min, followed by 30 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min at 72 °C. Following the last cycle, an elongation step of 10 min at 72 °C was applied. The PCR program for amplification of the *nifH* gene was according to the nested PCR protocol as described by Zani *et al.* (2000) with modifications. In brief, the first PCR using primers nifH 1 and 2 was started with a hot start of 15 min at 96 °C and followed by 35 cycles of 1 min at 94 °C, 1 min at 57 °C, 1 min at 72 °C, followed by 10 min of elongation at 72 °C. One microliter of the PCR product was then used in the second PCR using the primers nifH 3 and 4. The PCR program started with a 15 min hot start at 94 °C, followed by 35 cycles of 1 min at 94 °C, 1 min at 54 °C, 1 min at 72 °C, followed by 10 min of elongation at 72 °C.

The DNA clean & concentrator-5 kit (Zymogram, Zymo Research, Orange, CA, USA) was used to remove primer dimers from the PCR reactions following the instructions of the manufacturer. DNA concentration of the purified PCR products was checked spectrophotometrically (Nanodrop, ND1000). For sequencing reactions, 3.5 µl of the purified and concentrated PCR product served as template using 10 µM of the sequencing primer and the Big Dye Terminator v1.1 cycle sequencing kit (Applied Biosystems) following the manufacturer's instructions. Sequencing primers were the same forward and the reverse primers as were used for PCR amplification (Table 2). Sequence products were analyzed using a 3130 Genetic Analyzer (Applied Biosystems). Sequences were edited manually using ChromasPro V 1.41 (Technelysium Pty Ltd, Tewantia, Queensland, Australia) and manually checked for errors in base calling. Only high-quality sequences were included in the final data set. The sequences were deposited in GenBank under the following accession numbers: 16S rRNA (EU025781-EU025806), ITS-1 region (EU119301-EU119325), 23S-129F (EU025807-EU025831), 23S-2241R (EU025756-EU25780) and the *cpcBA* operon (EU119326-EU119352).

Phylogenetic analysis
All *Pseudanabaena* sequences obtained from this study were aligned using CLUSTALW integrated into the package BioEdit (Hall, 1999). These sequences were aligned against sequences closely related to *Pseudanabaena/Limnothrix* group previously identified by BLASTN comparison from GenBank and other references identified in previous published studies. Sequence comparison and phylogenetic analyses of the partial sequencing of 16S rRNA, 23S rRNA, ITS-1, *cpcBA* and intergenic spacer (IGS) were performed using the software MEGA3.1 (Kumar *et al.*, 2004). Neighbor-joining with Jukes–Cantor correction and 1000 bootstraps was used to build the corresponding phylogenetic trees. Partial amino-acid sequences of the coding region of the *cpcBA* operon were also used in the phylogenetic analyses performed with the neighbor-joining method as well as with maximum parsimony. Maximum parsimony was used with the close-neighbor-interchange search algorithm with random tree addition using 100 bootstraps.

Population genetic analysis
The protein coding nucleotide sequences from the *cpcBA* locus aligned by ClustalW were analyzed using DnaSP version 4.0. (Rozas *et al.*, 2003) to calculate the following parameters: (i) synonymous and nonsynonymous polymorphic sites, (ii) estimation of Ka/Ks divergence ratio, (iii) The McDonald–Kreitman test (McDonald and Kreitman, 1991) to detect positive or purifying selection and the HKA test (Hudson, Kreitman, Aguadé) on the basis of neutral theory of molecular evolution (Hudson *et al.*, 1987), (iv) degree of genetic differentiation between populations, estimated by F_{st} (Hudson *et al.*, 1992), (v) estimation of the recombination parameter (R) and minimum number of recombination events (Hudson *et al.*, 1987). Finally, Selecton version 2.2 (http://selecton.bioinfo.tau.ac.il) was used to identify positive and purifying selection at each of the amino acids using a Bayesian inference approach (Stern *et al.*, 2007). The Selecton server automatically calculates the ratio between Ka and Ks (ω) at each codon site using a maximum likelihood approach. The value of ω at each site is translated to a discrete color scale projected onto one of the homologous sequences for each sequence clusters. Colors 1 to 2 (dark and light yellow) indicate $\omega > 1$ and stand for sites with positive selection, whereas the shades of white through magenta (colors 3 through 7) indicated various levels of $\omega \leqslant 1$ where Selecton results can be accurately used to infer sites undergoing purifying selection.

Results

Phenotypic traits: morphological characteristics and photosynthetic pigment composition
A total of 28 strains were isolated from two geographically distant locations: AV, Spain (10 strains isolated in 1997) and the BS (1 strain isolated in 1995, 17 strains in 2004). The isolates were assigned to *Pseudanabaena* on the basis of their morphological characteristics, such as cell size, motility on agarose plates and the presence of polar

gas vesicles (Table 1). The isolates displayed different cell dimensions ranging from 2.1×1.9 to $6.7 \times 1.4\,\mu m$ (length × width) (Supplementary Figure 1SM). While cell width remained within a narrow range, cell length varied considerably. All strains, except CCY9710 from AV, possessed polar gas vesicles (Table 1). Gliding motility was observed in most of the strains except for four isolates from AV (Table 1).

The presence of the major light-harvesting pigments chlorophyll *a* (Chl*a*), PC and PE was determined by *in vivo* absorption spectra, from which the ratios PE:PC, Chl*a*:PE and Chl*a*:PC were calculated (Table 1). *Pseudanabaena* isolates were divided into strains that have both PE and PC pigments (63% of all strains) and those with only PC (37%) as their major pigment (Figure 1). All strains possessing both PE and PC pigments were capable of CCA. PC-rich strains reveal an absorption peak at ~625 nm and, hence, harvest orange-red light effectively. The proportion of strains capable of CCA that were isolated from the BS was 66%, slightly higher than the 55% for the AV. The majority of strains that were positive for CCA possessed a ratio of the absorption at 570 and 625 nm of approximately one or more when incubated in white light. When incubated under red or green light, cultures changed pigmentation toward green and red, respectively (Figures 1b and c, Supplementary Figure 2SM). Chlorophyll *a* showed absorption peaks at 440 nm (Soret band) and 680 nm. The absorption peaks of the three major light harvesting pigments were at the same wavelengths in all isolates but their relative heights varied substantially (Table 1).

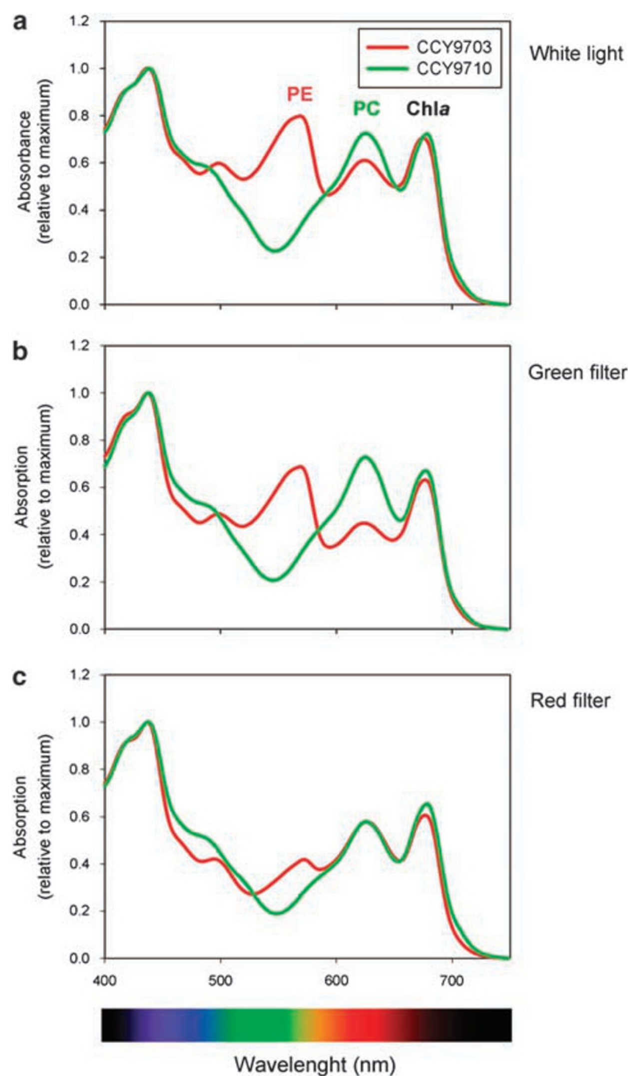
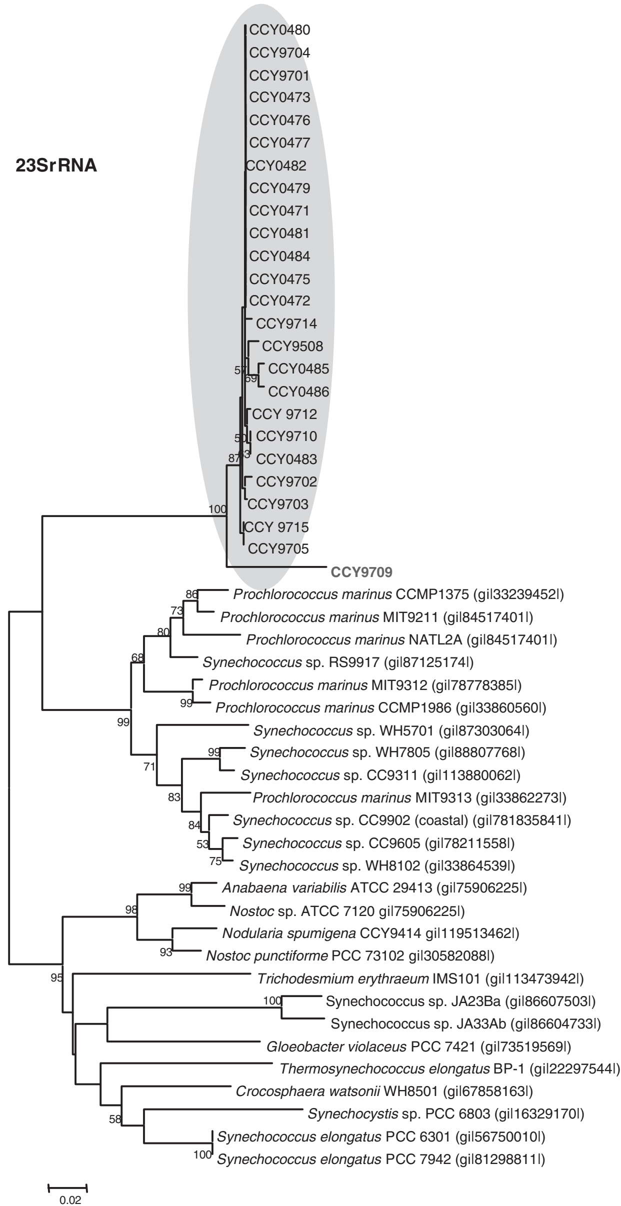


Figure 1 The effect of light colour on the pigment composition of *Pseudanabaena* strains CCY9703 and CCY9710. Three cultures of each strain were grown under white light until the cultures reached the stationary phase. The cultures were then transferred and grown under white (**a**), green (**b**) or red (**c**) light. Strain CCY9703 shows a decreased phycoerythrin (PE) absorption under the influence of red light and an increased phycocyanin (PC) absorption indicating its capacity for complementary chromatic adaptation (CCA). Strain CC9710 showed a decrease in PC absorption under influence of red light (PC-rich strain).

The phylogeny of Pseudanabaena *revealed from their 16- and 23S rRNA genes*

To determine the phylogenetic relationships of the *Pseudanabaena* isolates, the 16- and 23S rRNA genes were partly sequenced and analyzed. Because the 23S rRNA gene offers a higher phylogenetic resolution, the start and end of the 23S rRNA gene were both sequenced using the forward 129F and reverse 2241R primers, respectively (Table 2). This resulted in two products with a sequence length of 440 bp (129F) and 527 bp (2241R), respectively. The sequences obtained by using these primers gave similar phylogenetic relationships, although a larger number of polymorphisms were observed when using primer 129F (data not shown). Figure 2 depicts the phylogeny on the basis of neighbor-joining analysis and compares the 23S rRNA (Figure 2a) with the 16S rRNA gene tree topologies (Figure 2b) obtained using the primers 129F (440 bp) and Bact1055F (400 bp), respectively. In both trees, all isolates grouped into a single 99% similarity cluster indicating congruency in the phylogeny of the 16- and 23S rRNA genes. Moreover, the 23S rRNA gene displayed more polymorphisms than the 16S rRNA gene, which was virtually identical in virtually all isolates. The phylogenetic analysis of the 16S rRNA gene confirmed that all isolates belong to the *Pseudanabaena/Limnothrix* group with 99% cluster similarity (Zwart *et al.*, 2005; Willame *et al.*, 2006). From BLAST searches against the GenBank database, we observed that 42% of the isolates were 100% identical to *Pseudanabaena* sp. PCC6903 (AM709632). The other strains possess high similarity (99%) to *Pseudanabaena* sp. 1tu24s9 (AM259269), which originated from the Finnish freshwater Lake Tuusulanjarvi (Supplementary Table 1SM). Only strain CCY9709 from AV exhibited a higher divergence



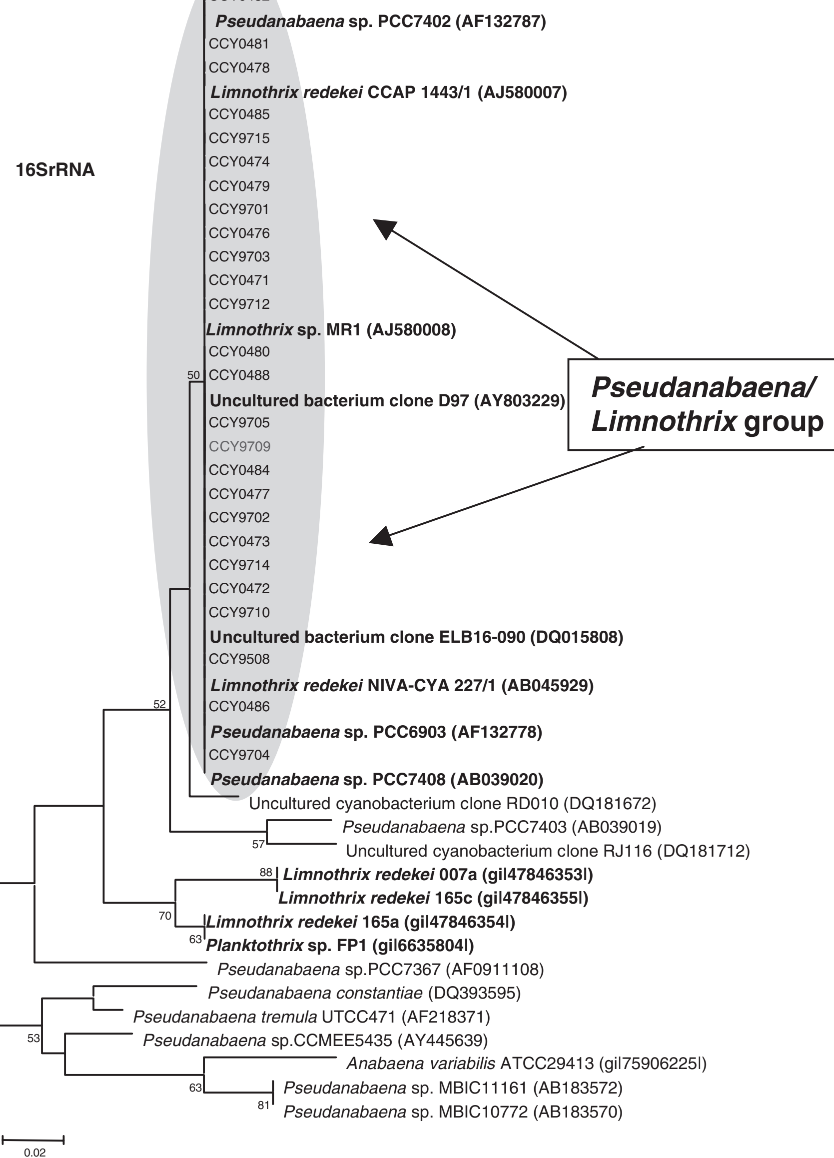


Figure 2 Neighbor-joining tree based on comparison of the partial sequences of the 23S rRNA (**a**; 129F primer) and 16S rRNA (**b**; Bact1055F primer) genes of the *Pseudanabaena* strains collected in this study. All *Pseudanabaena* strains grouped in a single 99% cluster, for both the 23- and the 16S rRNA genes (shadowed area).

on the basis of the 23S rRNA sequence (Figure 2a) showing a 94% similarity with *Pseudanabaena* sp. PCC6903 (Supplementary Table 1SM).

Microdiversity within Pseudanabaena *strains: ITS sequencing analysis*

The partial sequencing of the 16- and 23S rRNA genes was not sufficient to resolve the phylogeny of the *Pseudanabaena/Limnothrix* group. Therefore, the ITS region located between the 16- and 23S rRNA genes was sequenced. All *Pseudanabaena* ITS sequences revealed the same structure and contained two tRNA genes, tRNA*Ile* and tRNA*Ala* (data not shown). However, there was a high nucleotide divergence and length variability among the ITS sequences. The phylogenetic analysis of the ITS revealed a higher level of differentiation within the *Pseudanabaena/Limnothrix* group (Figure 3), which constrained most of the strains in three major 99% clusters ('microdiversity clusters'). Hence, ITS exhibited a higher degree of microdiversity compared to the 16S or 23S rRNA gene sequences where all sequences clustered together in one single 99% similarity cluster (Figure 2). Unfortunately, no ITS sequences related to *Pseudanabaena* species have been published to date, and therefore our sequences clustered with the two known *Limnothrix* ITS sequences.

The ITS analysis uncovered different patterns with clusters containing isolates specific for the BS, isolates specific for the AV, and a mixed cluster containing isolates from both locations. Most of the BS strains grouped in the BS cluster (BSC) with 99.6% similarity. BSC contained 12 isolates from the BS but also one strain from the AV (CCY9710) and *Limnothrix* sp. MR1 (isolated from Lake Loosdrecht, The Netherlands (Zwart *et al.*, 2005)). The mixed Albufera and BSC (MABSC) contained five identical sequences retrieved from isolates from both locations. The fact that strains originating from such distant locations and with 7 years between their isolation possess identical ITS sequences is at least surprising, since it is well-known that ITS is highly variable (in length and/or sequence) and therefore changes were expected. The AV (AVC) comprised six isolates from AV possessing 99.2% similarity. Strain CCY9709 grouped in a fourth cluster together with *Limnothrix redekei* Culture Collection of Algae and Protozoa (CCAP) 1443/1. Furthermore, unique

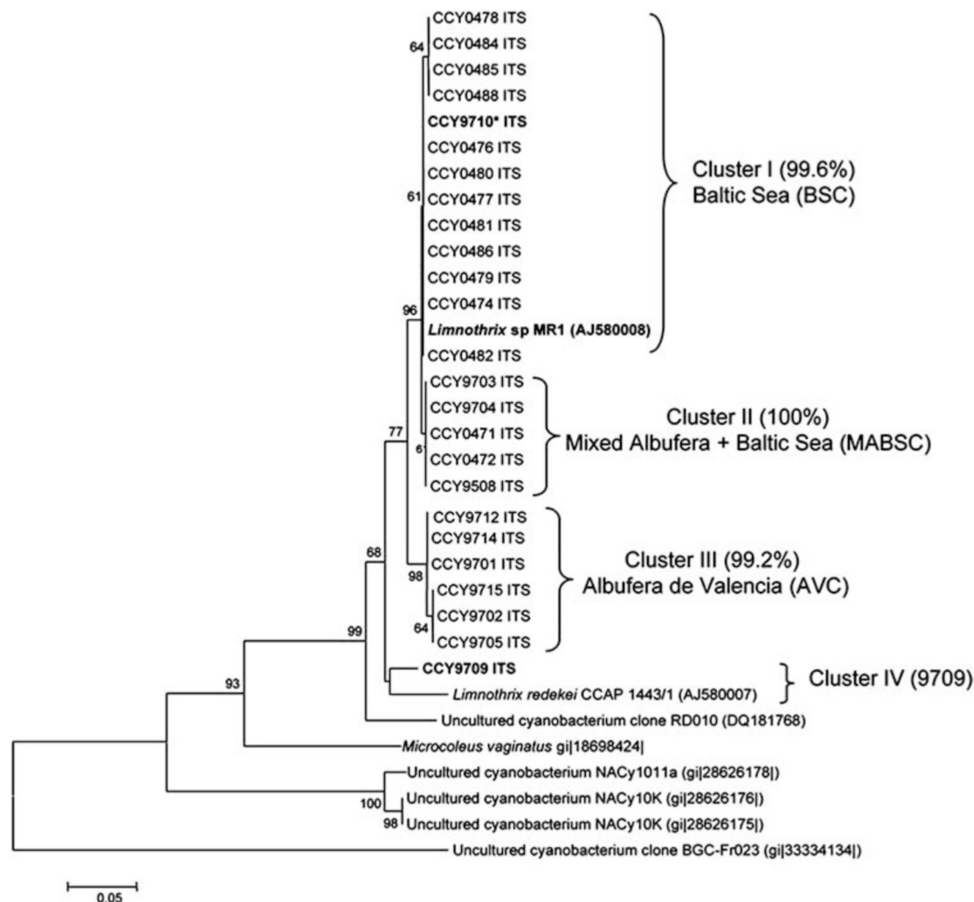


Figure 3 Microdiversity of *Pseudanabaena* spp. as revealed by ITS-1 analysis. The ITS allowed further differentiation showing geographical patterns, with clusters specific for the Baltic Sea (cluster I) and Albufera de Valencia, Spain (cluster III), but also a cluster with representatives of both locations (cluster II). Strain CCY9710 from the Albufera de Valencia represents the only exception and clustered with the Baltic Sea strains. ITS, internal transcribed spacer.

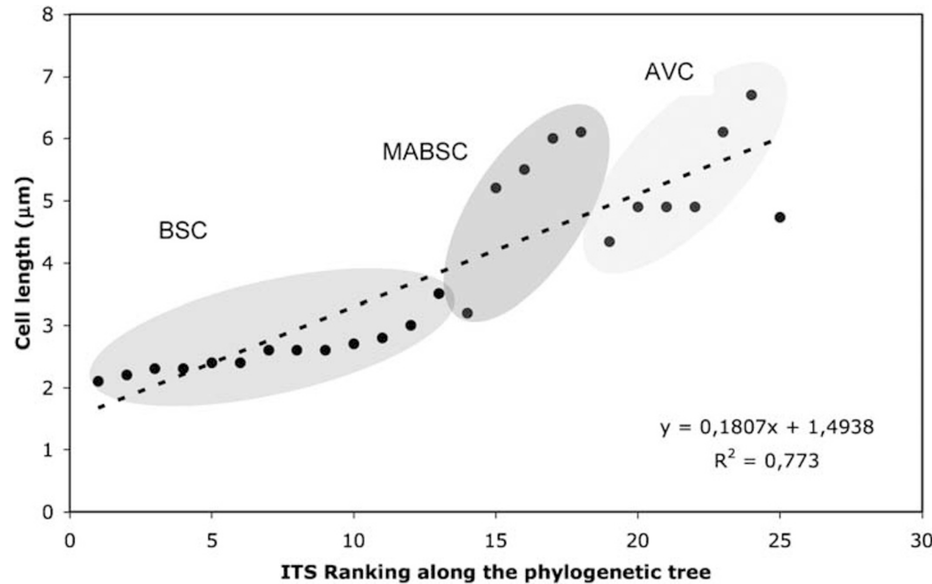


Figure 4 Correlation between cell length and the ranking of *Pseudanabaena* strains in the phylogenetic tree derived from the internal transcribed spacer 1(ITS-1) sequences. BSC correspond to the Baltic Sea cluster, MABSC to the mixed Albufera and BSCand AVC to the Albufera de Valencia cluster. The blue dot corresponds to strain CCY9709. See online version for color figure.

insertion sequences were found in some clusters. A unique 36 bp insertion sequence was found in the MABSC. In CCY9709, three specific insertion sequences of 23, 6 and 5 bp were found. The position of the isolates in the ITS phylogenetic tree showed a close relationship with cell length (Figure 4). The cell length of all BSC strains, including CCY9710 from AV, ranged from 2.1 to 3.5 µm. The MABSC cluster was characterized by cells that are slightly longer, ranging from 3.2 to 6.1 µm. Finally, the AV cluster comprised the strains with the longest cells, ranging from 4.3 to 6.7 µm. Strain CCY9709 possessed cells of 4.9 µm long and grouped in an independent monophyletic cluster with *L. redekei* CCAP 1443/1 which, however, has considerably longer cells (6–10 µm; JG Day, personal communication with the CCAP).

Phylogeny of the phycocyanin operon: correlation of the cpcBA *gene clusters with light absorption spectra*

A major limitation of the use of ribosomal genes as molecular markers is the impossibility to attribute ecophysiological traits to these genes. Therefore, we also sequenced the *cpcBA*. *CpcBA* encodes the two subunits of PC, which is part of the phycobilisome, the major light-harvesting complex of cyanobacteria. The main goal was to assign sequence clusters of *cpcBA* to the *in vivo* light absorption spectra of the isolated strains to reveal ecologically different populations (ecotypes) within *Pseudanabaena* (Figure 5). The *cpcBA* locus has been widely used for the study of cyanobacterial diversity and phylogeny, and it is therefore suitable for our purpose (Ivanikova *et al*., 2007; Six *et al*., 2007; Haverkamp *et al*., 2008). In addition, the IGS between *cpcB* and *cpcA* (*cpcBA*-IGS) was sequenced (Supplementary Figure 3SM and Figure 5b) to explore whether coding and noncoding regions of the *cpcBA* operon display different evolutionary rates resulting in different phylogenies.

Phylogenetic analysis of the partial sequences of *cpcBA* encoding for 150 amino acids, displayed two well-supported clusters (Figure 5a) with similarities higher than 99% ('microdiversity clusters'). Cluster I comprises sequences with 99.9% similarity and grouped 20 isolates from both locations. Half of these isolates were rich in PC (PC-rich) and unable to perform CCA, whereas the other 10 strains were capable of CCA. This cluster is closely related to *Pseudanabaena* sp. PCC7409 that is capable of CCA. Cluster II contained only strains capable of CCA, both from the BS and the AV and possess 100% sequence similarity. Half of the strains of cluster II investigated for *nifH* PCR amplification possessed *nifH* encoding for dinitrogenase reductase, which is a component of nitrogenase. Among cluster I, there was only one strain (CCY0477) that possessed *nifH* in a total of 20 isolates. *NifH* was found in strains originating from both environments. *CpcBA* cluster II in Figure 5 corresponds to the MABSC cluster of ITS sequences in Figure 4. This hints to the presence of a conserved and coherent lineage within *Pseudanabaena*. Again, the CCA positive strain CCY9709 displayed the most divergent position in the *cpcBA* phylogeny.

The phylogeny of the concatenated *cpcBA*-IGS sequences revealed geographical patterns that were not shown in the *cpcBA* phylogeny alone (Figure 5b). The tree topology of the concatenated *cpcBA*-IGS phylogeny was consistent with the topology of the *cpcBA* phylogeny without the IGS

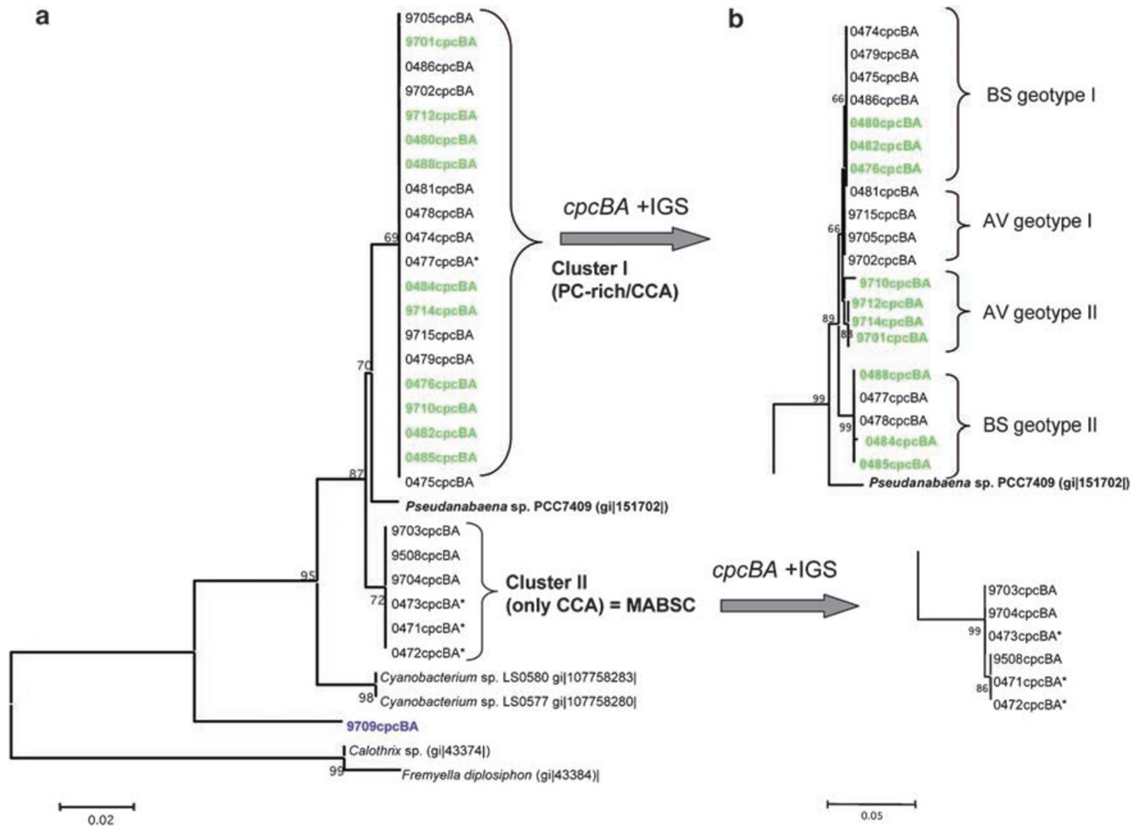


Figure 5 (**a**) Phylogenetic relationships of *Pseudanabaena* strains using 150 amino acids of the phycocyanin operon (*cpcBA*) genes. The green labeled strains represent PC-rich *Pseudanabaena* spp. All other strains contained both PC and phycoerythrin and are capable of complementary chromatic adaptation (CCA). Asterisks indicate *nifH*-positive strains. Cluster II consists entirely of CCA strains, and is identical to MABSC in the ITS-based phylogenetic tree (see Figure 3). (**b**) Phylogenetic relationships derived from the *cpcBA* operon concatenated with the intergenic spacer (IGS; about 600 bp). This revealed different subclusters ('geotypes') from the Baltic Sea (BS) and Albufera de Valencia (AV) within cluster I. ITS, internal transcribed spacer; MABSC, mixed Albufera and Baltic Sea cluster; PC, phycocyanin.

region. Moreover, the heterogeneity and length variability of the IGS sequences allowed a higher level of differentiation and revealed different subclusters from specific geographic locations ('geotypes'). Four geotypes were discerned within cluster I (PC-rich/CCA), two of them originated from the BS and the other two from Alfubera de Valencia (Figure 5b). In all geotypes, specific base-pair signatures were assigned (Supplementary Figure 3SM). Complete IGS sequence length ranged from 104 to 212 bp. The strains of cluster I (PC-rich/CCA) possessed an IGS of 153 bp, whereas the IGS sequences of cluster II (only CCA) were shorter and possessed only 104 bp. Strain CCY9709 had the longest (212 bp) and most divergent IGS (Supplementary Figure 3SM). Moreover, cluster II (only CCA) comprised different IGS sequences compared to cluster I-IGS and contained several deletions (in total 49 bp). Nonetheless, cluster II-IGS sequences were conserved and only a few specific signatures were found. For example, the T at position 120 or the C at position 132 are representative for two of the three BS strains that possess *nifH* (Supplementary Figure 3SM).

Evolutionary forces operating on the phycocyanin operon genes

On the basis of the analysis of the *cpcBA* operon, the evolutionary processes shaping the PC genes among *Pseudanabaena* spp. were investigated. Partial sequencing of 450 nucleotides constrained 26 strains in two major lineages, clusters I (20 isolates) and II (six isolates; Figure 5a). The number of polymorphisms within and between both lineages was examined. A total of 16 single nucleotide polymorphisms were found between both lineages and therefore the nucleotide divergence observed between both clusters was 3.82%. The number of synonymous (silent) nucleotide substitutions per synonymous site (Ks = 0.0809) was 12.3-fold higher than the nonsynonymous substitutions (amino-acid change) per nonsynonymous site (Ka = 0.00655). The ratio of nonsynonymous and synonymous nucleotide substitutions in protein-coding genes (Ka/Ks) may give important clues about the selection pressure on and evolution of the protein-coding genes (McDonald and Kreitman, 1991). The ratio of Ka/Ks observed for *cpc*BA was 0.074. Ratios of Ka/Ks < 1 indicate that purifying selection takes

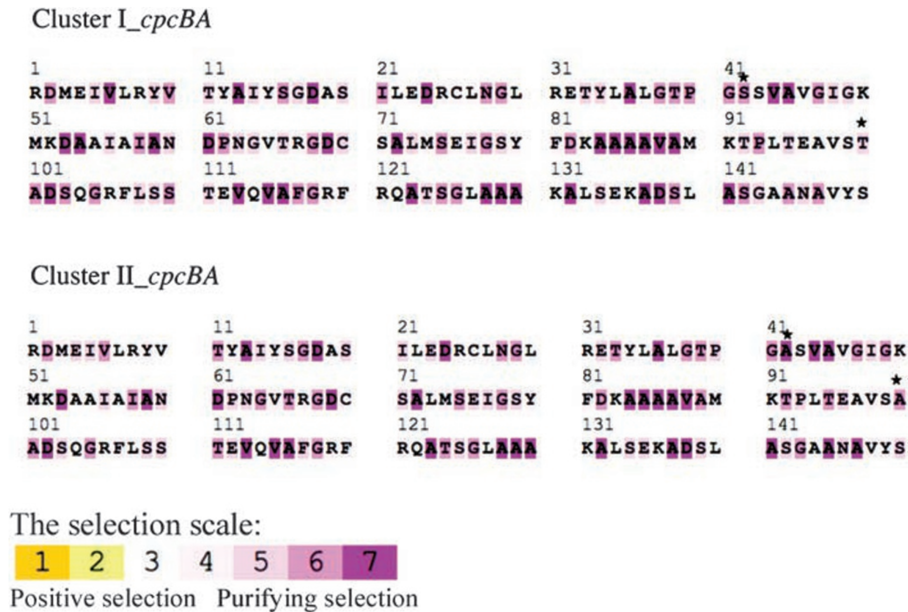


Figure 6 Selecton results for the *cpcBA* locus on 20 sequences from cluster I and 6 sequences from cluster II using the E8 model (Stern *et al*., 2007). Different levels of purifying selection are colored in shades of magenta through 150 amino acids of the *cpcBA* locus. The scale bar indicates selection ranging from positive (yellow colors 1 and 2) to purifying selection (colors 3 through 7). No positive selection is observed. The star symbols represent the nonsynonymous sites (amino-acid changes) between clusters I and II.

place. Hence, this is the case for the *cpcBA* clusters I and II. Purifying selection means that nonsynonymous vs synonymous mutations is favored in the direction of the latter. As a consequence, over time, slightly deleterious nonsynonymous mutations are continuously removed from the population leaving only synonymous mutations. In addition, the possibility to detect different types of selection (positive or purifying) on specific amino acids within the *cpcBA* operon was explored (Stern *et al*., 2007). Positive selection was not detected in any of the amino acids, which indicates that all amino acids must have been under different levels of purifying selection (Figure 6). Furthermore, the degree of genetic differentiation between both clusters (*Pseudanabaena* subpopulations) was also explored by estimating the fixation index F_{st}, which indicates the amount of gene flow between populations (Hudson *et al*., 1992). Values of F_{st} range between 0 and 1. The value of 0 indicates that the populations share the same alleles, whereas the value of 1 shows that the populations are fixed for different alleles. The F_{st} value for the *cpcBA* loci was 0.886, which indicated high gene flow compared with other described bacterial populations (Whitaker *et al*., 2003; Miller *et al*., 2006).

Finally, the number of amino-acid changes (nonsynonymous sites) was assessed for the 150 codons analyzed. Only two amino-acid changes were detected, yielding a 1.4% amino-acid divergence. In both cases, the identity of the amino acids differed systematically between clusters I and II of the *Pseudanabaena cpcBA* lineages. At position 42 at the end of the beta subunit (*cpcB*), a serine (S, alcohol polar R-group) found in cluster I was exchanged in cluster II by an alanine (A, aliphatic R-group, a hydrophobic molecule). Moreover, at position 100, at the beginning of the α-subunit (*cpcA*) a threonine (T, a hydrophilic and hydroxyl-containing amino acid) in cluster I was replaced by alanine (A) in cluster II. However, neither of these two amino acids were found to be under positive selection (white color in Figure 6).

Discussion

Pseudanabaena/Limnothrix *group: agreement of phenotype with genetic data*

Pseudanabaena is morphologically similar to *Limnothrix*. The large variation of cell lengths observed among the strains isolated in this study suggests that *Pseudanabaena* shows a high level of plasticity. Accordingly, cell length is not very useful for distinguishing *Pseudanabaena* from *Limnothrix*.

On the basis of 16S rRNA gene sequences, earlier studies showed that *Pseudanabaena* and some strains of *Limnothrix* cluster with *Pseudanabaena* spp., including one isolate assigned to *L. redekei*. This cluster is commonly referred to as the *Pseudanabaena/Limnothrix* group (Gkelis *et al*., 2005; Willame *et al*., 2006). This group also comprised several environmental sequences as well as other isolates including the type strain *Pseudanabaena* PCC7408, strains belonging to *L. redekei* (Van Goor, Meffert) (*Limnothrix* sp. MR1 from Lake Loosdrecht, *L. redekei* CCAP 1443/1 and *L. redekei* CCAP 227/1) and several other isolates from Lake Loosdrecht, The Netherlands (Zwart *et al*., 2005; Willame *et al*., 2006). A second cluster, comprising only *L. redekei*

strains isolated from Lake Kastoria (Greece), has also been observed (Gkelis *et al.*, 2005; Willame *et al.*, 2006). On the basis of partial sequencing of the 16- and 23S rRNA genes, all *Pseudanabaena* strains are isolated in this study cluster with the *Pseudanabaena/Limnothrix* group. Neither 16- nor 23S rRNA gene analysis could distinguish between *Pseudanabaena* and *Limnothrix*, confirming previous reports (Zwart *et al.*, 2005; Willame *et al.*, 2006). However, ITS analysis of the *Pseudanabaena* isolates shows a high nucleotide divergence and size variability revealing a higher level of differentiation in this group. For instance, strains belonging to the AVC and strain CCY9709 all possessed cells that were almost 3 times longer than wide (Figure 4), a characteristic that is considered typical for *Limnothrix*. This could indicate that these strains are indeed closely related to *L. redekei*. Nonetheless, in order to confirm these observations, it is necessary to increase the number of ITS sequences of *L. redekei* and *Pseudanabaena*. Moreover, it will be necessary to re-evaluate the morphological basis on which the genera *Limnothrix* and *Pseudanabaena* are separated.

Distribution and abundance of Pseudanabaena

The distribution of *Pseudanabaena* is widespread, although their actual abundance is still to be determined. Microscopic observations confirm their occurrence as a major fraction of the summer cyanobacterial blooms in the BS and AV. Remarkably, environmental clone libraries from the BS targeting the 16S rRNA–ITS and the *cpcBA* and *cpeBA* genes encoding the phycocyanin and phycoerythrin operons, respectively did not recover any *Pseudanabaena* sequences (Haverkamp *et al.*, 2008). However, it is a well-known fact that the environmental sequences are strikingly different from the organisms isolated from most microbial communities (Suzuki *et al.*, 1997; Eilers *et al.*, 2000, 2001; Stevens *et al.*, 2005; Donachie *et al.*, 2007). We cannot exclude the possibility that *Pseudanabaena* filaments were retained by the 20 μm nylon mesh filter that was used as a prefiltration step and that this was the reason that we did not recover their sequences. Nevertheless, *Pseudanabaena* strains were isolated from the <20 μm fraction. Moreover, Denaturing Gradient Gel Electrophoresis (DGGE) fingerprinting using cyanobacteria-specific 16S rRNA primers detected *Pseudanabaena* in the BS (B Díez, Stockholm University, personal communication and unpublished results). As we used general bacterial primers for the construction of our environmental clone libraries, we may have missed out *Pseudanabaena* because of the far more abundant unicellular picocyanobacteria.

Members of *Pseudanabaena* have been reported from a variety of different environments including freshwater lakes (Gkelis *et al.*, 2005; Zwart *et al.*, 2005; Kim *et al.*, 2006; Willame *et al.*, 2006), brackish environments (Stal *et al.*, 2003), hot springs (Castenholz *et al.*, 2001) as well as from epilithic cyanobacterial communities of beach rock (Heron Island, Great Barrier Reef) (Diez *et al.*, 2007). Our study revealed several lineages of *Pseudanabaena* from the BS and from the AV, on the basis of a multilocus sequence typing approach using five loci of 28 isolates. The average genetic divergence of each of the markers varied from 1% in the small and large subunit ribosomal RNA genes to 3.5% in ITS-1 and 3.82% in *cpcBA*. The BS is geographically distant from the AV. Moreover, they represent quite different habitats (i.e., brackish vs freshwater) and it is, therefore, not surprising to find different *Pseudanabaena* genotypes at these two locations. Endemic clusters (geotypes) were detected in the BSC and AVC on the basis of the analysis of ITS sequences (Figure 3). The analysis of the *cpcBA*-IGS operon revealed an even better differentiation of these geotypes (Figures 5b, Supplementary Figure 3SM).

Yet, the BS is connected to the North Sea, and through the Atlantic Ocean and the Mediterranean Sea, it ultimately links to AV. This connection could allow the long-range dispersal of *Pseudanabaena* genotypes, although other mechanisms for long-range dispersal could also be envisioned (such as transport by birds). Salinity differences between the BS and the AV are minor, and organisms might adapt quickly to somewhat higher or lower salt levels. Indeed, a coherent and 'cosmopolitan' monophyletic cluster, comprising members from both locations was found for several loci (MABSC by ITS and cluster II by *cpcBA* genes or *cpcBA*-IGS analysis) suggesting a conserved and homogeneous lineage within *Pseudanabaena* (Figures 3 and 5). This lineage clustered the strains isolated from the two distant locations, even though the isolation at the two locations was separated by 7 years. This could point to the global dispersal of these *Pseudanabaena* strains. However, more *Pseudanabaena* strains from other environments and locations should be included to confirm the possibility of global dispersal. On the basis of the analysis of the genes of the PC operon and the light absorption spectra associated with them, it is proposed that this lineage or subpopulation of *Pseudanabaena* (MABSC by ITS and cluster II by *cpcBA* genes or *cpcBA*-IGS analysis) represents an ecotype that possesses the ability of CCA. Interestingly, this ecotype also possesses the *nifH* gene in at least half of their members, and therefore might be capable of N_2 fixation (Figure 5). This ecotype proposed with the name of Pseudanabaena Complementary Chromatic Adaptation (PCCA) is probably widely distributed, and the ability of CCA might provide it with a selective advantage.

Genetic diversification of Pseudanabaena *populations*

Correlation of the *cpcBA* gene clusters with the light absorption spectra hinted at the coexistence of two *Pseudanabaena* populations with a niche differen-

tiation along the light spectrum (Figure 5, clusters I and II). Cluster II or the 'PCCA ecotype' occurs as a group with 99% similarity ('microdiversity cluster'), at a variety of different loci (ITS, *cpcBA* gene and *cpcBA*-IGS analysis). Previous studies indicate that such microdiversity clusters could represent important units of differentiation as ecotypes in natural populations of bacteria (Palys *et al*., 1997; Moore *et al*., 1998; Rocap *et al*., 2003; Konstantinidis and Tiedje, 2005; Lopez-Lopez *et al*., 2005; Thompson *et al*., 2005; Cohan, 2006; Polz *et al*., 2006; Cohan and Perry, 2007), and are often observed in environmental clone libraries (Field *et al*., 1997; Acinas *et al*., 2004; Morris *et al*., 2005; Johnson *et al*., 2006; Pommier *et al*., 2007). The microdiversity clusters identified here are correlated with morphological and ecophysiological traits such as cell length and the capacity to perform CCA (Figures 4 and 5), providing further support for the designation as an ecotype (Ahlgren and Rocap, 2006; Johnson *et al*., 2006; Polz *et al*., 2006; Ward *et al*., 2006). Although this genetic pattern agrees with the ecotype model for bacterial species (Palys *et al*., 1997; Cohan, 2002; Gevers *et al*., 2005; Cohan and Perry, 2007), other mechanisms causing the genetic diversification of *Pseudanabaena* populations cannot be excluded. Indeed, the loss of substructure in the tree topology occurred when different loci were compared (e.g., ITS vs *cpcBA*-IGS). For instance, strain CCY9710 from AV fell into the BSC on the basis of its ITS sequence and it shared a similar cell length with all strains. However, this strain grouped with other isolates from AV when considering the phylogeny on the basis of *cpcBA*-IGS sequences. Similarly, the MABSC was an independent branch in the phylogeny of the *cpcBA* locus, whereas it was sister to the BSC using ITS. Moreover, at least 12 recombination events (R_m) were detected at the *cpcBA* locus, emphasizing the importance of homologous recombination and that this process should be taken into account.

By using the ratio of nonsynonymous vs synonymous fixation as a measure of the level of selective pressure on the *cpcBA*, it was concluded that purifying selection is involved in the evolutionary diversification of *Pseudanabaena* populations. Other population genetic analyses such Tajima's D and the Mc Donald–Kreitman (MK) tests for selection were not significant (data not shown), further supporting our finding that the mutations at *cpcBA* do not deviate from those expected from neutrality and, hence, are not under positive selection. The results show that divergence of the *cpcBA* in clusters I and II is promoted by purifying selection in both populations. Evidence for purifying selection in cyanobacteria exist, for instance, for *hetR* of *Trichodesmium*, *nifH* of *Cylindrospermopsis raceborskii* and *rpoC1* of *Anabaena lemmermannii* (Mes and Stal, 2005; Mes *et al*., 2006). Our results extend this list and provide the first report on evolutionary diversification in *Pseudanabaena* genus confirming similar evolutionary trends as described previously for other cyanobacterial taxa. However, this should not be generalized as other studies detected positive selection for PE genes of *Prochlorococcus* and *Synechococcus* (Qin *et al*., 2005; Zhao and Qin, 2007) and for other functional genes such *kaiC* of *Microcoleus chthonoplastes* and *rbcX* of *Anabaena* and *Aphanizomenon* sp. (Mes *et al*., 2006). Moreover, it is also known that *hetR* of different lineages of *Trichodesmium* (Mes and Stal, 2005) or the *kai* genes, family of *Nostoc linckia,* (Dvornyk *et al*., 2002) undergo different selective forces and that similar genes like *sasA* and *kai* genes in *Synechococcus* are subject to different selective constrains (Dvornyk *et al*., 2004).

Recent laboratory experiments investigated the role of CCA in the competition of *Pseudanabaena* against red and green *Synechococcus* strains (Stomp *et al*., 2008). The competition experiments showed that *Pseudanabaena* was a strong competitor in fluctuating light environments, provided that it had sufficient time to adjust its pigment composition to the prevailing light spectrum. *Pseudanabaena* can change its pigmentation from red to green, and *vice versa*, within $\sim$7 days. Thus, *Pseudanabaena* benefited from CCA only if fluctuations in underwater light color were slow compared with the time required for CCA, corresponding to slow mixing processes or infrequent storms in their natural habitat (Stomp *et al*., 2008). We hypothesize that PC-rich strains in cluster I have lost the capacity of CCA recently by the loss or presence of dysfunctional genes required to synthesize the PE disks in the phycobilisome, or the genes like *rcaE* which is needed for the control of CCA (Terauchi *et al*., 2004; Kehoe and Gutu, 2006). Knockout experiments targeting the *rcaE* gene showed that this gene is needed for responsiveness to both red and green light under CCA (Terauchi *et al*., 2004). The loss or the presence of a dysfunctional copy of *rcaE* or other genes involved in the phycobilisomes or CCA may be caused by deleterious mutations, resulting in strains that have lost PE and are only able to use PC and Chl*a* as their main light harvesting pigments. PC absorbs photons in the orange-red part of the light spectrum. Accordingly, loss of CCA is likely to be advantageous in moderately turbid waters where orange-red light predominates (Stomp *et al*., 2007), and during storm periods with rapid mixing when CCA is too slow to track changes in the underwater light spectrum experienced by the entrained *Pseudanabaena* filaments (Stomp *et al*., 2008). Variation in the underwater light spectrum at a range of different time scales could thus have induced genetic divergence between PC-rich and CCA strains because of differences in fitness. The proposed mechanism of selective sweeps could enable such genetic diversification (Cohan, 2002, 2006; Gevers *et al*., 2005; Cohan and Perry, 2007).

Analyses of other functional genes such as those coding for PE will provide further insights into the

mechanisms of diversification within *Pseudanabaena*. Also, it would be interesting to explore with competition experiments the fitness adaptation to recognize any correlation between these two traits and establish their ecological implications.

In summary, multilocus sequencing of five independent loci revealed the existence of several lineages or subpopulations within *Pseudanabaena*. The phylogenies of the 16- and 23S rRNA genes are consistent, but analysis of the other loci indicated loss of substructure, suggesting the recombination between these loci. *Pseudanabaena* isolates exhibited high levels of microdiversity unveiling different patterns with both local as well as more globally dispersed populations. A conserved *Pseudanabaena* lineage proposed as PCCA ecotype was characterized by the capacity of chromatic adaptation and possibility for N_2 fixation. Population genetic analyses of the PC genes suggest an evolutionary diversification of *Pseudanabaena* through purifying selection.

The isolation of additional *Pseudanabaena/Limnothrix* strains from a variety of different environments is required to further elucidate of the ecology, biogeography and evolution of this understudied group of bloom-forming cyanobacteria.

Acknowledgements

We are grateful to Marije Doeleman for her assistance in sequencing and to Ute Wollenzien for isolation, culturing and maintenance of the *Pseudanabaena* strains. We thank Maayke Stomp and the crew of the RV Aranda for their assistance during sampling of the Baltic Sea. Special thanks to Caroline Jenkins for providing us the 0.45 μm filters during the Cyano-cruise in 2004 and Dr. Beatríz Díez for sharing her unpublished data with us. Research of JH was supported by the Earth and Life Sciences Foundation (ALW), which was subsidized by the Netherlands Organization for Scientific Research (NWO). SGA was supported in the framework of the Flemish-Dutch collaboration for Sea Research (VLANEZO), TH and LJS acknowledge the support from the European Commission through the project MIRACLE (EVK3-CT-2002–00087). This is publication 4365 of NIOO-KNAW.

References

Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL *et al.* (2004). Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**: 551–554.

Ahlgren NA, Rocap G. (2006). Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. *Appl Environ Microbiol* **72**: 7193–7204.

Biegala IC, Not F, Vaulot D, Simon N. (2003). Quantitative assessment of picoeukaryotes in the natural environment by using taxon-specific oligonucleotide probes in association with tyramide signal amplification-fluorescence *in situ* hybridization and flow cytometry. *Appl Environ Microbiol* **69**: 5519–5529.

Castenholz RW, Rippka R, Herdman M, Wilmotte A.. (2001). Form-genus XII. Pseudanabaena Lauterborn 1916. In: Boone DR, Castenholz RW (eds). *Bergey's Manual of Systematic Bacteriology*, 2nd edn, Springer Verlag: Heidelberg, pp 554–557.

Chorus I, Bartram J. (1999). *Toxic Cyanobacteria in Water*. Taylor & Francis: London, UK.

Cohan FM.. (2002). What are bacterial species? *Annu Rev Microbiol* **56**: 457–487.

Cohan FM.. (2006). Towards a conceptual and operational union of bacterial systematics, ecology, and evolution. *Philos T Roy Soc B* **361**: 1985–1996.

Cohan FM, Perry EB. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Curr Biol* **17**: R373–R386.

Diez B, Bauer K, Bergman B. (2007). Epilithic cyanobacterial communities of a marine tropical beach rock (Heron Island, Great Barrier Reef): Diversity and diazotrophy. *Appl Environ Microbiol* **73**: 3656–3668.

Donachie SP, Foster JS, Brown MV. (2007). Culture clash: Challenging the dogma of microbial diversity. *ISME J* **1**: 97–99.

Dvornyk V, Deng HW, Nevo E. (2004). Structure and molecular phylogeny of *sasA* genes in cyanobacteria: insights into evolution of the prokaryotic circadian system. *Mol Biol Evol* **21**: 1468–1476.

Dvornyk V, Vinogradova O, Nevo E. (2002). Long-term microclimatic stress causes rapid adaptive radiation of *kaiABC* clock gene family in a cyanobacterium, *Nostoc linckia*, from 'Evolution Canyons' I and II, Israel. *Proc Natl Acad Sci USA* **99**: 2082–2087.

Eilers H, Pernthaler J, Glockner FO, Amann R. (2000). Culturability and *in situ* abundance of pelagic bacteria from the North Sea. *Appl Environ Microbiol* **66**: 3044–3051.

Eilers H, Pernthaler J, Peplies J, Glockner FO, Gerdts G, Amann R. (2001). Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton. *Appl Environ Microbiol* **67**: 5134–5142.

Ernst A, Becker S, Wollenzien UIA, Postius C. (2003). Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. *Microbiol SGM* **149**: 217–228.

Field KG, Gordon D, Wright T, Rappe M, Urbach E, Vergin K,*et al.* (1997). Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Appl Environ Microbiol* **63**: 63–70.

Garcia-Martinez J, Acinas SG, Anton AI, Rodriguez-Valera F. (1999). Use of the 16S-23S ribosomal genes spacer region in studies of prokaryotic diversity. *J Microbiol Methods* **36**: 55–64.

Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ *et al.* (2005). Re-evaluating prokaryotic species. *Nat Rev Microbiol* **3**: 733–739.

Gkelis S, Rajaniemi P, Vardaka E, Moustaka-Gouni M, Lanaras T, Sivonen K. (2005). *Limnothrix redekei* (Van Goor) Meffert (Cyanobacteria) strains from Lake Kastoria, Greece form a separate phylogenetic group. *Microb Ecol* **49**: 176–182.

Granéli E, Turner J. (2006). *Ecology of Harmful Marine Algae*. Springer: Berlin, Germany, pp 1–413.

Hall TA. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.

Haverkamp T, Acinas SG, Doeleman M, Stomp M, Huisman J, Stal LJ. (2008). Diversity and phylogeny of Baltic Sea picocyanobacteria inferred from their ITS and phycobiliprotein operons. *Environ Microbiol* **10**: 174–188.

Hudson RR, Kreitman M, Aguade M. (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153–159.

Hudson RR, Slatkin M, Maddison WP. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**: 583–589.

Huisman J, Matthijs HCP, Visser PM. (2005). *Harmful Cyanobacteria*. Springer: Berlin, Germany, pp 1–241.

Hunt DE, Klepac-Ceraj V, Acinas SG, Gautier C, Bertilsson S, Polz MF. (2006). Evaluation of 23S rRNA PCR primers for use in phylogenetic studies of bacterial diversity. *Appl Environ Microbiol* **72**: 2221–2225.

Ivanikova NV, Popels LC, McKay RML, Bullerjahn GS. (2007). Lake Superior supports novel clusters of cyanobacterial picoplankton. *Appl Environ Microbiol* **73**: 4055–4065.

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, Chisholm SW. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**: 1737–1740.

Kehoe DM, Gutu A. (2006). Responding to color: the regulation of complementary chromatic adaptation. *Annu Rev Plant Biol* **57**: 127–150.

Kim SG, Rhee SK, Ahn CY, Ko SR, Choi GG, Bae JW *et al.* (2006). Determination of cyanobacterial diversity during algal blooms in Daechung Reservoir, Korea, on the basis of *cpcBA* intergenic spacer region analysis. *Appl Environ Microbiol* **72**: 3252–3258.

Komárek J. (2003). Planktic oscillatorialean cyanoprokaryotes (short review according to combined phenotype and molecular aspects). *Hydrobiologia* **502**: 367–382.

Konstantinidis KT, Tiedje JM. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA* **102**: 2567–2572.

Kumar S, Tamura K, Nei M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**: 150–163.

Lane DJ. (1991). 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds). *Nucleic Acid Techniques in Bacterial Systematics*. John Wiley and Sons Ltd, New York, 115–147.

Lopez-Lopez A, Bartual SG, Stal L, Onyshchenko O, Rodriguez-Valera F. (2005). Genetic analysis of housekeeping genes reveals a deep-sea ecotype of *Alteromonas macleodii* in the Mediterranean Sea. *Environ Microbiol* **7**: 649–659.

Mayer J, Dokulil MT, Salbrechter M, Berger M, Posch T, Pfister G,*et al.* (1997). Seasonal successions and trophic relations between phytoplankton, zooplankton, ciliate and bacteria in a hypertrophic shallow lake in Vienna, Austria. *Hydrobiologia* **342**: 165–174.

McDonald JH, Kreitman M. (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.

Mes THM, Stal LJ. (2005). Variable selection pressures across lineages in *Trichodesmium* and related cyanobacteria based on the heterocyst differentiation protein gene *hetR*. *Gene* **346**: 163–171.

Mes THM, Doeleman M, Lodders N, Nubel U, Stal LJ. (2006). Selection on protein-coding genes of natural cyanobacterial populations. *Environ Microbiol* **8**: 1534–1543.

Miller SR, Purugganan MD, Curtis SE. (2006). Molecular population genetics and phenotypic diversification of two populations of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Appl Environ Microbiol* **72**: 2793–2800.

Moore LR, Rocap G, Chisholm SW. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**: 464–467.

Morris RM, Vergin KL, Cho JC, Rappe MS, Carlson CA, Giovannoni SJ. (2005). Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. *Limnol Oceanogr* **50**: 1687–1696.

Palys T, Nakamura LK, Cohan FM. (1997). Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int J Syst Bacteriol* **47**: 1145–1156.

Polz MF, Hunt DE, Preheim SP, Weinreich DM. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos T Roy Soc B* **361**: 2009–2021.

Pommier T, Canback B, Riemann L, Bostrom KH, Simu K, Lundberg P *et al.* (2007). Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* **16**: 867–880.

Qin S, Zhao FQ, Tseng CK. (2005). Evidence for positive selection in phycoerythrin genes of red algae and cyanobacteria *Prochlorococcus* and *Synechococcus*. *Photosynthetica* **43**: 141–146.

Rippka R, Deruelles J, Waterbury JB. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* **111**: 1–61.

Rippka R, , Herdman M. (1992). *Pasteur Culture Collection of Cyanobacteria*, vol. 1. Institute Pasteur: Paris, France.

Robertson BR, Tezuka N, Watanabe MM.. (2001). Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. *Int J Syst Evol Microbiol* **51**: 861–871.

Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, Ahlgren NA *et al.* (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042–1047.

Romo S, Miracle MR. (1994). Population dynamics and ecology of subdominant phytoplankton species in a shallow hypertrophic lake (Albufera of Valencia, Spain). *Hydrobiologia* **273**: 37–56.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Rücker J, Wiedner C, Zippel P. (1997). Factors controlling the dominance of *Planktothrix agardhii* and *Limnothrix redekei* in eutrophic shallow lakes. *Hydrobiologia* **342**: 107–115.

Six C, Thomas JC, Garczarek L, Ostrowski M, Dufresne A, Blot N *et al.* (2007). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol* **8**: R259.

Stal LJ, Albertano P, Bergman B, von Brockel K, Gallon JR, Hayes PK *et al.* (2003). BASIC: Baltic Sea cyanobacteria. An investigation of the structure and dynamics of water blooms of cyanobacteria in the Baltic Sea—responses to a changing environment. *Cont Shelf Res* **23**: 1695–1714.

Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. (2007). Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res* **35**: W506–W511.

Stevens H, Stubner M, Simon M, Brinkhoff T. (2005). Phylogeny of Proteobacteria and Bacteroidetes from oxic habitats of a tidal flat ecosystem. *FEMS Microbiol Ecol* **54**: 351–365.

Stomp M, Huisman J, de Jongh F, Veraart AJ, Gerla D, Rijkeboer M *et al.* (2004). Adaptive divergence in pigment composition promotes phytoplankton biodiversity. *Nature* **432**: 104–107.

Stomp M, Huisman J, Vörös L, Pick FR, Laamanen M, Haverkamp T *et al.* (2007). Colourful coexistence of red and green picocyanobacteria in lakes and seas. *Ecol Lett* **10**: 290–298.

Stomp M, van Dijk MA, van Overzee HMJ, Wortel M, Sigon CAM, Egas M *et al.* (2008). The time scale of phenotypic plasticity, and its impact on competition in fluctuating environments. *Am Nat* (in press).

Suzuki MT, Rappe MS, Haimberger ZW, Winfield H, Adair N, Strobel J *et al.* (1997). Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. *Appl Environ Microbiol* **63**: 983–989.

Terauchi K, Montgomery BL, Grossman AR, Lagarias JC, Kehoe DM. (2004). RcaE is a complementary chromatic adaptation photoreceptor required for green and red light responsiveness. *Mol Microbiol* **51**: 567–577.

Thompson JR, Pacocha S, Pharino C, Klepac-Ceraj V, Hunt DE, Benoit J *et al.* (2005). Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**: 1311–1313.

Vasconcelos VM, Pereira E. (2001). Cyanobacteria diversity and toxicity in a wastewater treatment plant (Portugal). *Water Res* **35**: 1354–1357.

Villena MJ, Romo S. (2003). Phytoplankton changes in a shallow Mediterranean lake (Albufera of Valencia, Spain) after sewage diversion. *Hydrobiologia* **506**: 281–287.

Ward DM, Bateson MM, Ferris MJ, Kuhl M, Wieland A, Koeppel A *et al.* (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos T Roy Soc B* **361**: 1997–2008.

Whitaker RJ, Grogan DW, Taylor JW. (2003). Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**: 976–978.

Willame R, Boutte C, Grubisic S, Wilmotte A, Komarek J, Hoffmann L. (2006). Morphological and molecular characterization of planktonic cyanobacteria from Belgium and Luxembourg. *J Phycol* **42**: 1312–1332.

Zani S, Mellon MT, Collier JL, Zehr JP. (2000). Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. *Appl Environ Microbiol* **66**: 3119–3124.

Zhao F, Qin S. (2007). Comparative molecular population genetics of phycoerythrin locus in *Prochlorococcus*. *Genetica* **129**: 291–299.

Zwart G, Kamst-van Agterveld MP, van der Werff-Staverman I, Hagen F, Hoogveld HL, Gons HJ. (2005). Molecular characterization of cyanobacterial diversity in a shallow eutrophic lake. *Environ Microbiol* **7**: 365–377.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)